Volume 6 • Number 1 • 2003
VALUE IN HEALTH

Computerized Health Information and the Demand for Medical Care

Todd H. Wagner, PhD,[1] Holly B. Jimison, PhD[2]

[1]VA Health Economics Resource Center, Palo Alto Medical Center, Palo Alto, CA, USA, and the Department of Health Research and Policy, Stanford University School of Medicine, Stanford, CA, USA; [2]Kaiser Permanente Center for Health Research, Portland, OR, USA

ABSTRACT

Objective: Consumer health information, once the domain of books and booklets, has become increasingly digitized and available on the Internet. This study assessed the effect of using computerized health information on consumers' demand for medical care.

Methods: The dependent variable was self-reported number of visits to the doctor in the past year. The key independent variable was the use of computerized health information, which was treated as endogenous. We tested the effect of using computerized health information on physician visits using ordinary least squares, instrumental variables, fixed effects, and fixed-effects instrumental variables models. The instrumental variables included exposure to the Healthwise Communities Project, a community-wide health information intervention; computer ownership; and Internet access. Random households in three cities were mailed questionnaires before and after the Healthwise Communities Project. In total, 5909 surveys were collected for a response rate of 54%.

Results: In both the bivariate and the multivariate analyses, the use of computerized health information was not associated with self-reported entry into care or number of visits. The instrumental variables models also found no differences, with the exception that the probability of entering care was significantly greater with the two-stage conditional logit model ($P < .05$).

Conclusions: Although providing people with health information is intuitively appealing, we found little evidence of an association between using a computer for health information and self-reported medical visits in the past year. This study used overall self-reported utilizations as the dependent variable, and more research is needed to determine whether health information affects the health production function in other important ways, such as the location of care, the timing of getting care, or the intensity of treatment.

Keywords: medical utilization, information, interactive health communication, self-care information, Internet.

Introduction

In 1997, the National Library of Medicine (NLM) stopped charging for access to MEDLINE. The occasion created much fanfare and at the unveiling, as then Vice President Gore said, "This development, by itself, may do more to reform and improve the quality of health care in the United States than anything else we have done in a long time" [1].

Since then, it is hard to imagine that anyone could have predicted how quickly the landscape would have changed. The NLM continues to provide free access to MEDLINE. But consumer health information, once the domain of books and booklets, has become increasingly digitized and available on the Internet. Major health organizations, such as the American Medical Association and Kaiser Permanente, have invested heavily in proprietary health information systems. In addition, smaller start-up companies have pushed the envelope on the World Wide Web. Consumers can now access everything from risk assessment tools, to interactive health advice, to the latest medical news on the Internet.

Yet, what is the value of this information? Until recently, a strong capital market and the desire to push the technological envelope overshadowed this question. In 1998 and 1999, approximately \$3 billion in private and public equity was invested in "e-health" companies [2]. By May 2000, the stocks at many of these companies were down 60% or more from their 52-week highs [2]. Wall Street uses short-term benchmarks to gauge success, and many

Address correspondence to: Todd H. Wagner, VA Health Economics Resource Center, Stanford University School of Medicine, 795 Willow Road (MPD 152), Menlo Park, CA 94025. E-mail: twagner@stanford.edu

 1098-3015/03/$15.00/29 29–39

of these companies became mired with the complexities of medicine. Although many investors may find other industries more profitable in the short term, it is still valuable for the medical community to understand more about the role of computerized health information.

Many expect that only positive things will come from providing consumers with more health information. Yet, some have suggested that more health information may lead to worsening outcomes from "information overload" [3]. Searching, obtaining, and processing health information take time and cognitive effort [4]. When ill, most people are under time constraints and stressed, making it more difficult to understand the information. This situation is made worse by the fact that much of the information on the Internet is inaccurate [5,6].

If in fact consumer health information is beneficial, is there any supporting evidence? A common assumption is that health information, specifically self-care information, helps people reduce the unnecessary use of medical care. Some clinical trials have found significant reductions in utilization when people are provided with various forms of health information [7–9]. At the same time, other studies have found either no effect or the opposite effect [10,11]. A recent evidence-based report reviewed these clinical trials and concluded that despite the mixed findings, there is reason to believe that self-care information reduces the demand for medical care [12]. The evidence-based report evaluated clinical trials and did not review community-based studies. Kenkel [13] and Hsieh and Lin [14] used national surveys of the United States and Taiwan, respectively, to assess the effects of information on the demand for medical care. These studies found that more informed and more knowledgeable patients seek preventive care and favor health behaviors that improve health. Although consistent with the clinical trials, these studies only assessed the use of health information through indirect proxies. They did not have any direct measures of health information use. Therefore, they measured an individual's level of information by summing up correct responses to health-related questions. This assumes that health knowledge is related to health information use, which may be inappropriate. As Borgmann [15] reminds us, information should not be confused with knowledge. To date, direct measures of information utilization have been lacking in population-based studies. In addition, many different types of health and self-care information exist, such as the Internet, telephone advice nurses, and health reference books. Most studies have provided people with specific information, and little has been done to evaluate the effect of the Internet.

This paper uses data from a community-based informational intervention. The intervention was designed to provide information through three sources: self-care books, telephone advice nurses, and computers and Web pages. Elsewhere we have assessed the use of health information from the intervention [16] and the intervention's effect on self-reported utilization and pediatric visits [17,18]. In those papers, we used an intent-to-treat analysis to compare respondents in the intervention city to respondents in the control cities. We did not distinguish between the types of information the person used. Given the interest in computerized health information, this study uses a different methodologic approach that compares users of computerized health information to nonusers. Consistent with past work, we treat the use of information as a choice variable, that is, endogenous, and use instrumental variables for identification. The new contribution of this study is that the data come from a community-wide intervention and we know from self-report whether the individual used a computer for health information in the past few months.

Methods

The Intervention: Healthwise Communities Project

Late in the summer of 1996, Healthwise Incorporated started the Healthwise Communities Project (HCP) in its home city, Boise, Idaho. The community-wide intervention was designed to provide all Boise area residents with health information and self-care resources. Each Boise household, consisting of a community population of 132,000, was sent the *Healthwise Handbook*, which is a self-care and health reference guide. For the Boise community, a toll-free health advice nurse was established. To enable residents to use computers to access health information, information stations, which included computers connected to information databases, books, and other health consumer information, were set up in public libraries, businesses, and health care settings throughout the Boise area. Boise residents could also access the health database on the Internet by entering their zip code. Although not strictly designed to keep out non-Boise residents, it is unlikely that many, if any, residents in the control cities used the database. Finally, Healthwise Inc. sponsored workshops to help Boise residents and physicians develop self-care and health-related communication skills. As part of the entire program, Healthwise initiated a large, multimedia advertising campaign. Previous research indicates that the use of self-care resources associated with the HCP was significant [16].

Conceptual Model and Hypotheses

The use of information can affect one's utilization of medical care through two primary pathways. The first and simplest pathway is that people use these new health information resources instead of getting health information from their doctor. This is a substitution hypothesis and elsewhere we have reported some evidence suggesting that this occurred with the HCP [19]. The second pathway is an efficiency hypothesis, whereby a person who uses health information as an input into Grossman's [20] health production model and becomes a more efficient consumer of health care. Ideally, smart consumers would be able to triage their symptoms, treating minor problems and seeking care when appropriate. Some research unrelated to the HCP has found that health information, specifically health knowledge, is related to the use of medical care [14,21]. Unlike the substitution framework, this pathway is more complex because it depends on people's cognitive abilities to process the information and to disentangle necessary care from unnecessary care. While we can suggest alternative links between using health information and the utilization of health care, it is not clear whether the two pathways will have complementary or negating effects.

Ideally, if we were to study the effects of the HCP, we would have different measures of utilization, including emergency room and urgent care visits, preventive care, and inpatient stays. Health information probably affects these types of utilization in different ways. However, because of practical constraints, the data set under investigation only asked about overall number of visits to the doctor.

Data

As part of the evaluation of the HCP, 7500 surveys (2500 per site) were mailed to randomly selected households in Boise, Idaho; Eugene, Oregon; and Billings, Montana. This baseline survey was carried out in the spring of 1996, before the start of the HCP. The households of Eugene and Billings were chosen as controls based on geographical proximity to Boise, metropolitan characteristics, and similarities in health systems. A list of householder names for all three sites was purchased from a national marketing agency. A simple random sample was then drawn from the list. The survey and postage-paid return envelope were sent to the householder. Of the 7500 baseline surveys distributed, 1048 were returned as undeliverable and 1 as deceased. In total, 3067 surveys were completed. Of the delivered surveys, 47.5% responded. In January 1998, those who responded to the baseline survey were sent a follow-up survey. Because of concerns about attrition, a second random sample of 3600 households (1200 per site) was sent a follow-up survey at the same time. The same survey was used for all persons in 1998. Of the 6667 persons sent a follow-up questionnaire, 2090 were returned as undeliverable, 12 were marked deceased, and 2842 were completed. Of the surveys delivered, 62% were completed. Overall, 5909 surveys were completed, representing an adjusted response rate of 54%. Although lower than some of the national phone health surveys, these response rates are consistent or higher than past postal studies that did not use financial incentives [22,23].

To assess medical utilization, the survey asked, "In the last 12 months, about how many times have you visited a doctor (not including eye doctors or dentists)?" Respondents were then asked to select one of 6 responses: 0 visits, 1 visit, 2 to 3 visits, 4 to 5 visits, 6 to 10 visits, and 11 or more visits. To turn these categorical responses into an ordinal scale, each category was assigned the following values: 0 visits, 1 visit, 3 visits, 5 visits, 8 visits, and 16 visits, respectively. Approximately 5% of the sample had 11 or more visits. They were assigned 16 visits as an average. This was based on a Commonwealth Fund National Survey [24], which showed that approximately 5% of the people had 16 or more visits. In subsequent analyses, the assignment of an average of 16 was varied between 11 and 30; all conclusions remained the same. With this variable, we compared entry into care (used any care vs. did not use any care) and number of visits. Number of visits was not conditional upon using any care; thus we did not use the commonly used two-part model. The two-part model presents problems when following people over time as some nonusers at baseline used care at follow-up and vice versa. Also, other concerns about the two-part model exist [25].

In addition to analyzing number of visits, we also evaluated entry into care. Prior research has shown that imperfect information leads people to underestimate the marginal product of medical care [26]. Therefore, we expected the HCP intervention to increase the probability of entry into care. We also predicted that the intervention would reduce the number of visits. This prediction was based on a previous finding that people use self-care information to help determine when medical care is needed and then use it as a substitute for professional care and advice when appropriate [12]. Consequently, providing a person with health information might encourage more appropriate utilization, and the effects on entry into care and volume of care may work in opposite directions.

Independent variables included age, household income, educational attainment, insurance status, gender, employment status, health status, a list of 10 chronic conditions, having one or more children, marital status, travel time to the doctor, computer ownership, and access to the Internet. Variable names, definitions, and means are listed in Table 1.

Table 1 Percentages of the dependent and independent variables

	Full sample			Panel*	
	Total (N = 5909)	Control (n = 4010)	Exp. (n = 1899)	Control (n = 2660)	Exp. (n = 1110)
Dependent variables					
Number of visits: 0–16 visits (mean)	4.06	4.06	4.06	4.08	4.11
Entry into care: 0 = no; 1 = yes	86.4	86.4	86.6	86.3	86.8
Independent variables					
Year: 0 = 1996; 1 = 1998	48.1	48.6	47.0	50.0	50.0
INT: 0 = control; 1 = experiment	32.1	0.0	100.0	0.0	100.0
Used computer health info: 0 = no; 1 = yes	11.8	11.2	13.1	11.2	14.6
Income					
<$15,000	17.7	19.6	13.7	19.4	13.7
$15,000–$24,999	32.4	33.6	29.8	33.1	27.8
$25,000–$49,999	20.1	19.6	21.2	19.7	21.9
$50,000+	29.8	27.2	35.3	27.9	36.7
Education					
High school	30.5	33.2	24.8	34.2	23.6
Some college	33.1	33.0	33.5	32.1	33.1
College graduate	20.9	20.3	22.1	19.6	21.1
Postgraduate work	15.5	13.6	19.6	14.1	22.3
Insurance					
None	9.4	9.7	8.8	7.1	6.8
Any private	78.7	76.7	82.7	79.7	84.2
Only public	11.9	13.5	8.5	13.3	9.0
Sex: 0 = male; 1 = female	68.7	68.4	69.3	69.2	70.4
Race: 0 = not white; 1 = white	96.7	96.7	96.8	97.4	98.0
Age (years)					
18–29	9.7	8.9	11.7	4.9	6.5
30–44	29.5	27.0	34.7	25.2	32.9
45–64	37.4	38.1	35.7	41.1	39.3
65+	23.5	26.1	17.6	28.8	21.4
High blood pressure	22.0	22.6	20.6	24.4	22.9
High cholesterol	19.9	21.1	17.2	22.4	19.5
Arthritis	21.2	21.1	21.5	23.0	25.2
Chronic back pain	13.6	13.3	14.2	14.3	15.6
Cancer	3.8	3.9	3.7	4.3	4.0
Heart disease	6.0	6.6	4.8	7.3	4.9
Diabetes	4.7	4.9	4.4	4.7	3.9
Depression	12.3	12.0	13.0	11.9	12.4
Asthma	7.2	7.3	7.0	7.0	6.7
Chronic bronchitis	3.3	3.3	3.2	3.7	2.9
Health status					
Excellent	15.7	14.7	17.8	15.2	16.9
Very good	36.6	36.5	36.8	34.8	36.9
Good	34.2	34.4	33.9	35.9	35.7
Fair or poor	13.5	14.4	11.5	14.1	10.5
Employ					
Working full or part time	57.7	54.7	63.8	52.3	60.7
Retired or homemaker	37.5	40.3	31.7	43.8	35.8
Unemployed or student	4.8	5.0	4.5	4.0	3.5
Has children: 0 = no; 1 = yes	32.2	29.8	37.4	26.2	34.1
Married: 0 = no; 1 = yes	68.6	68.0	69.9	68.9	72.1
Rural: 0 = urban; 1 = rural	19.5	20.1	18.3	22.6	20.9
Travel time (min) to MD					
<15	57.6	58.0	56.8	57.4	57.5
15–30	35.3	35.1	35.8	35.8	34.7
>30	7.1	6.9	7.4	6.9	7.8

Note: Numbers may not add up because of rounding.
*Panel included the subset of people followed over time.

Although no one variable had more than 15% missing data, a sufficient number of cases had some variables with missing data. Most statistical techniques drop cases with any missing data, also known as listwise deletion. Estimating the models with listwise deletion resulted in a loss of almost 3000 (approximately 50%) cases. Listwise deletion can lead to biased coefficients and a sizable loss of statistical efficiency [27]. We used the findings of Amelia et al. [28] to replicate the original data set five times. In each data set, Amelia et al. preserved the observed values and imputed the missing data using the EM algorithm, which is described in detail by Schafer [27]. Each data set is generated independently, and the missing values vary across the five data sets. Statistical analysis was carried out on each of the five "complete" data sets, and the results were then combined [29]. It is reassuring that analyzing the data with listwise deletion yielded very similar results, suggesting that the missing data were missing at random.

Analysis

The analysis involved regressing the dependent variable (use of medical care) on the use of computerized health information, health-care price factors (insurance and travel time), and demographic factors. The key variable was the use of computerized health information, which was assumed to be endogenous as using health information represents a choice.

We have different options for handling the endogeneity. First, we could include as many observable characteristics as possible in the multivariate models. Second, we could use instrumental variables (IV) regression to address the endogeneity. We have three potential instruments: computer ownership, access to the Internet, and exposure to the HCP citywide intervention. The two-stage instrumental variables approach for medical utilization is denoted as:

$$\begin{aligned}\text{info}_{it} &= \alpha_1 year + \alpha_2 INT + \alpha_3 INT \times year \\ &\quad + \alpha_4 P_i^{\text{info}} + \alpha_5 P_i^{MC} + \sum_j \alpha_j X_{ij} + v_{it}\end{aligned} \tag{1}$$

$$\begin{aligned}\text{utilization}_{it} &= \beta_1 \overline{\text{info}} + \beta_2 year + \beta_3 site \\ &\quad + \beta_4 P_i^{MC} + \sum_j \beta_j X_{ij} + \varepsilon_{it}\end{aligned} \tag{2}$$

where info_{it} is the use of computerized health information in the past few months. Dummy variables INT and year and the INT × year interaction identify the HCP intervention. The intervention, described above, made it easier for Boise residents without computers to use the Internet for health. In addition, a Web page was established to help Boise residents find and sort through health information. P^{info} represents whether people own a computer and whether they have access to the Internet at home or work. Demographic factors are abbreviated by X. P_i^{MC} represents the price of medical care, for which we use insurance status and travel time to the doctor as proxies. The predicted use of health information from Equation 1 is $\overline{\text{info}_{it}}$. The random error items are ν_{it} and ε_{it}.

The third option for handling the endogeneity involves analyzing the subset of persons followed over time. If the "health-nut effect" is time invariant, a fixed-effects model will remove the endogeneity. However, even with fixed effects, using computerized information may still be partially endogenous, and the fixed-effects model requires strict exogeneity [30]. To address this assumption, fixed effects can also be used with instrumental variables in a fixed-effects instrumental variables (FE IV) estimator. The FE IV model can be expressed as modeling the use of health information and then using the predicted value from Equation 3 in the structural model, Equation 4. The two-stage model for medical utilization can be expressed as:

$$\begin{aligned}(\text{info}_{i2} - \text{info}_{i1}) &= \alpha_1(year_2 - year_1) \\ &\quad + \alpha_2 INT(year_2 - year_1) + \Delta\alpha_4 P_i^{\text{info}} \\ &\quad + \Delta\alpha_5 P_i^{MC} + \Delta\sum_j \alpha_j X_{ij} + v_i\end{aligned} \tag{3}$$

$$\begin{aligned}(\text{utilization}_{i2} &- \text{utilization}_{i1}) = \\ &\beta_1 \overline{(\text{info}_2 - \text{info}_1)} + \Delta\beta_2 P_i^{MC} \\ &+ \Delta\beta_3 year + \Delta\sum_j \beta_j X_{ij} + \varepsilon_i\end{aligned} \tag{4}$$

The intuition behind the FE IV model is equivalent to the IV model. In practice, however, one needs to be careful because fixed effects can remove much of the variation, leaving instruments that are only weakly associated with the use of information. As has been shown, weak instruments can be problematic [31,32].

For entry into care, we estimated the following four regression models: probit, IV probit, conditional logit, and a two-stage conditional logit. The standard errors in the two-stage conditional logit model were corrected with bootstrapping based on 1000 replications. We also tried a logistic and IV logit model, but this had little effect on the results. For volume of care, we estimated an ordinary least squares regression, an IV regression, a fixed-effect regression, and a two-stage least squares model with fixed effects.

Instruments

Identification for the instrumental variables was obtained through exclusion restrictions. In the first stage, where the dependent variable was the use of information, we included all exogenous variables and proxies for the price of health information, that is, the HCP intervention effect, ownership of a computer, and access to the Internet. The HCP intervention effect was treated as a price because the HCP made it easier for Boise area residents to access computerized health information. These variables were then excluded when we modeled the use of medical care.

Using the HCP intervention effect as an instrument may be problematic because the HCP encouraged use of books, advice nurses, and computers. Previous research has shown that the HCP had a larger effect on the books than on the advice nurse or computers [16]. Therefore, this instrument may be proxy for using any of these health information resources. This may be less of a concern given the correlations between using a computer, books, and telephone advice nurses were small (< .20), but they were significant. Therefore, as a sensitivity analysis we re-estimated all of the instrumental variables models without the intervention effect.

Computer ownership and Internet access are valid instruments as long as they are not directly associated with using medical care or, more broadly, health status. To test this assumption, we ran two analyses. First, because there were more instruments than were necessary for identification, we ran overidentification tests [33]. The overidentification tests were negative. Second, we wanted to test whether the medical utilization or health status was associated with purchasing a computer. For this, we estimated conditional logit models using the cohort of people followed over time. We excluded people in the intervention site to minimize any spillover effects from the intervention. The conditional logit models showed that changes in time, employment status, and having children were associated with buying a computer and gaining access to the Internet. In no case was medical utilization, chronic condition, or health status statistically associated with getting a computer or Internet access (results not shown). High blood pressure was marginally significantly associated with Internet access ($p = .08$), and this may reflect an association owing to making multiple comparisons.

To check the strength of the instruments of intervention effect, computer ownership, and Internet access, we regressed the use of computerized information on the instruments. Without fixed effects, the intervention effect, computer ownership, and Internet access were strongly associated with using a computer ($t = 2.11$, $t = 8.78$, and $t = 11.88$, respectively). With fixed effects, the intervention effect ($t = 0.57$) and computer ownership ($t = 0.92$) were not strong instruments, whereas Internet access remained associated with using a computer ($t = 5.09$, respectively).

Results

Of the residents in the 3 Pacific Northwest communities, approximately 86% reported visiting a doctor in the last year (Table 1). The average number of visits was 4.06 (median 3). Estimates of self-reported utilization were very similar among the control and intervention sites for both the entire sample and the cohort of people followed over time.

Overall, 11.8% of the people reported using a computer for health information, as seen in Table 1. The use of computers changed over time. In 1996, only 10.5% of the people reported using a computer for health information in the past few months. By 1998, this percentage grew to 14.6%. This statistic is not surprising given that the number of households buying computers has grown because computer prices dropped precipitously during this time [34]. The rates of using computers for health information were slightly higher in Boise (intervention site) compared to the control sites ($p = .04$).

When we compared users of computerized health information to nonusers, the unadjusted data showed no significant differences either in entry into care or in volume of care, as shown in Table 2. As

Table 2 The effect of using computerized health information on the demand for medical care

		Computerized health information		
Self-reported utilization	Total, N	User (SD)	Nonuser (SD)	Difference
Number of visits (0–16 visits)	5117	4.04 (4.22)	4.06 (4.11)	−0.02
Entry (0-1)	5909	0.88 (0.33)	0.86 (0.34)	0.02

Note: No differences were statistically significant at $P < .05$.

Table 3 indicates, little changed when we adjusted for the variables that are listed in Table 1. In the instrumental variables models, users of computerized health information used slightly less care and were slightly more likely to have entered care; however, these effects were not significant. The primary difference that was apparent from analyzing the subset of people followed over time, that is, the panel, was the change in coefficient's sign (Table 4). The fixed-effect models indicated that there was a small, albeit insignificant, decrease in volume of care and a nonsignificant increase in entry into care. The FE IV models showed the same pattern: a nonsignificant decrease in volume of care and an increase in the likelihood of entering into care. The probability of entering care was significant, however, with the two-stage conditional logit model.

As a sensitivity analysis, we re-estimated the instrumental variables models excluding the HCP instrumental variable. This had very little effect on the results. Information use was not statistically associated with the volume of care. Again, in the two-stage conditional logit model, the use of computerized information was associated with a higher odds of enter into care ($p < .05$).

Conclusions

This study is one of the first to consider the use of computerized health information on the demand for medical care. These data show that the use of computerized health information was not associated with volume of care or entry into care, with the exception of the two-stage conditional logit model where using computerized information was associated with a higher probability of entering care.

At first pass, these findings contradict the two population-based studies of Kenkel [26] and Hsieh and Lin [14]. However, those studies used a much broader definition of health information that encompassed other forms of health information, such as books, and health knowledge. In comparison, this study only considered people's self-reported use of computerized health information in the past few months. Therefore, this study is really the first of its kind and it is difficult to directly compare this study to past research.

While the study had a sample of 5909 respondents, one concern may be that our sample size was too small. A post hoc power analysis shows that this study had more than 90% power to detect a difference of 0.5 visits. Although it is possible that the real effects were smaller and that the sample sizes were too small to detect them, it is unclear whether smaller effects would be clinically important. It should also be noted that the study's ability to detect a difference is related to the strength of the instruments [31,32]. Our three instruments were the intervention effect, owning a computer, and access to the Internet. In the non-fixed-effects models, these instruments were strongly associated

Table 3 Assessing the use of computerized health information on the demand for medical care: full sample

	Conditional models		Instrumental variables	
	Visits	Entry	Visits	Entry
Used computer health information	0.151	0.044	0.079	0.274
	(0.94)	(0.62)	(0.15)	(1.37)
Year: 0 = 1996; 1 = 1998	–0.021	0.121*	–0.015	0.100†
	(0.21)	(2.71)	(0.13)	(2.05)
INT: 0 = control; 1 = experiment	–0.018	0.017	–0.018	0.017
	(0.17)	(0.35)	(0.17)	(0.35)
Income				
$15,000–$24,999	0.144	0.195*	0.144	0.194*
	(0.90)	(2.88)	(0.90)	(2.86)
$25,000–$49,999	0.152	0.333*	0.154	0.323*
	(0.81)	(4.13)	(0.82)	(3.99)
$50,000+	0.167	0.461*	0.172	0.443*
	(0.87)	(5.53)	(0.88)	(5.22)
Education				
Some college	0.243	0.101	0.244	0.096
	(1.88)	(1.80)	(1.88)	(1.72)
College graduate	0.110	0.102	0.113	0.089
	(0.72)	(1.58)	(0.73)	(1.35)
Postgraduate work	0.496*	0.353*	0.500*	0.334*
	(2.95)	(4.59)	(2.92)	(4.26)

(continued)

Table 3 continued

	Conditional models		Instrumental variables	
	Visits	Entry	Visits	Entry
Insurance				
Any private	1.138*	0.424*	1.142*	0.419*
	(6.17)	(6.05)	(6.12)	(5.99)
Only public	1.327*	0.343*	1.329*	0.340*
	(5.77)	(3.61)	(5.77)	(3.59)
Sex: 0 = male; 1 = female	0.708*	0.377*	0.708*	0.378*
	(6.35)	(8.02)	(6.34)	(8.04)
Race: 0 = not white; 1 = white	–0.019	0.349*	–0.020	0.351*
	(0.07)	(3.22)	(0.07)	(3.25)
Age (years)				
30–44	–1.031*	–0.288*	–1.031*	–0.287*
	(5.36)	(3.60)	(5.36)	(3.60)
45–64	–0.825*	–0.236*	–0.827*	–0.227*
	(4.22)	(2.86)	(4.22)	(2.75)
65+	–1.255*	–0.206†	–1.260*	–0.191
	(5.33)	(2.02)	(5.30)	(1.85)
High blood pressure	0.401*	0.227*	0.400*	0.230*
	(2.96)	(3.40)	(2.94)	(3.45)
High cholesterol	0.122	0.158*	0.122	0.157*
	(0.92)	(2.43)	(0.92)	(2.41)
Arthritis	0.669*	0.293*	0.670*	0.291*
	(4.78)	(4.15)	(4.78)	(4.13)
Back pain	0.481*	0.156†	0.482*	0.153†
	(3.09)	(2.00)	(3.09)	(1.97)
Cancer	2.277*	0.743*	2.277*	0.741*
	(8.67)	(3.88)	(8.67)	(3.87)
Heart disease	1.015*	0.410*	1.013*	0.415*
	(4.46)	(3.05)	(4.45)	(3.10)
Diabetes	0.971*	0.235	0.971*	0.234
	(3.90)	(1.64)	(3.89)	(1.64)
Depression	1.784*	0.492*	1.786*	0.483*
	(11.18)	(5.47)	(11.14)	(5.35)
Asthma	0.633*	0.037	0.633*	0.037
	(3.17)	(0.39)	(3.17)	(0.39)
Chronic bronchitis	1.014*	0.225	1.014*	0.221
	(3.49)	(1.35)	(3.49)	(1.33)
Health status				
Very good	0.457*	0.209*	0.457*	0.210*
	(3.00)	(3.55)	(2.99)	(3.57)
Good	1.383*	0.466*	1.383*	0.465*
	(8.39)	(6.90)	(8.39)	(6.91)
Fair or poor	3.204*	0.647*	3.204*	0.645*
	(14.70)	(6.27)	(14.70)	(6.25)
Employ				
Retired or homemaker	0.076	0.030	0.076	0.032
	(0.56)	(0.51)	(0.56)	(0.55)
Unemployed or student	0.711*	0.008	0.712*	0.001
	(2.91)	(0.07)	(2.92)	(0.01)
Has children: 0 = no; 1 = yes	0.443*	0.008	0.444*	0.004
	(3.43)	(0.14)	(3.43)	(0.08)
Married: 0 = no; 1 = yes	–0.036	–0.025	–0.034	–0.032
	(0.29)	(0.47)	(0.27)	(0.60)
Rural: 0 = urban; 1 = rural	–0.048	0.006	–0.049	0.007
	(0.35)	(0.10)	(0.36)	(0.12)
Travel time to MD (min)				
15–30	0.058	0.073	0.057	0.073
	(0.53)	(1.51)	(0.52)	(1.52)
>30	0.336	–0.000	0.338	0.002
	(1.60)	(0.01)	(1.61)	(0.02)
Constant	1.041*	–0.589*	1.041*	–0.592*
	(2.58)	(3.71)	(2.58)	(3.74)
Observations	5909	5909	5909	5909
R^2	.18		.18	

Note: The difference is unadjusted. Absolute values of *t* statistics are in parentheses.
*Significant at 1% (two-tailed test).
†Significant at 5% (two-tailed test).

Table 4 Fixed-effects models testing the use of computerized health information on the demand for medical care: cohort followed over time

	Fixed effects		Fixed-effects instrumental variables	
	Visits	Entry	Visits	Entry*
Used computer health information	−0.194	1.036	−0.760	0.631†
	(0.52)	(1.47)	(0.52)	(2.75)
Year: 0 = 1996; 1 = 1998	−0.016	0.252	0.027	−0.066
	(0.11)	(1.18)	(0.15)	(0.22)
Income				
$15,000–$24,999	0.150	−0.175	0.117	0.315
	(0.30)	(0.22)	(0.23)	(0.38)
$25,000–$49,999	−0.343	0.149	−0.393	0.629
	(0.53)	(0.14)	(0.59)	(0.56)
$50,000+	−1.525†	−0.182	−1.622†	0.160
	(2.06)	(0.15)	(2.10)	(0.13)
Insurance				
Any private	0.031	0.775	−0.004	0.626
	(0.04)	(1.08)	(0.01)	(0.93)
Only public	−0.170	−2.590	−0.202	−0.777
	(0.21)	(1.81)	(0.24)	(0.82)
High blood pressure	0.823	1.536	0.826	1.126
	(1.71)	(1.60)	(1.71)	(1.23)
High cholesterol	0.858†	0.634	0.847†	0.345
	(2.22)	(0.85)	(2.18)	(0.49)
Arthritis	−0.109	0.255	−0.107	0.367
	(0.29)	(0.32)	(0.29)	(0.54)
Back pain	−0.008	0.158	−0.018	−0.013
	(0.02)	(0.19)	(0.04)	(0.02)
Depression	0.878	2.060†	0.870	1.906†
	(1.91)	(2.04)	(1.89)	(2.07)
Health status				
Very good	1.273‡	0.177	1.314‡	0.017
	(3.30)	(0.42)	(3.29)	(0.04)
Good	2.066‡	0.008	2.105‡	−0.402
	(4.44)	(0.01)	(4.42)	(0.58)
Fair or poor	2.612‡	−0.651	2.648‡	−1.031
	(4.01)	(0.64)	(4.01)	(1.08)
Employ				
Retired or homemaker	0.734	0.275	0.730	−0.184
	(1.60)	(0.45)	(1.58)	(0.32)
Unemployed or student	0.883	0.270	0.875	0.073
	(1.19)	(0.27)	(1.15)	(0.07)
Has children: 0 = no; 1 = yes	0.546	2.236	0.510	2.423†
	(0.98)	(1.86)	(0.90)	(2.02)
Married: 0 = no; 1 = yes	0.352	0.919	0.426	0.574
	(0.63)	(1.07)	(0.74)	(0.69)
Travel time to MD (min)				
15–30	0.381	0.295	0.364	0.319
	(1.30)	(0.77)	(1.23)	(0.88)
>30+	1.214	−0.439	1.228	−0.166
	(1.85)	(0.45)	(1.86)	(0.18)
Constant	1.699		1.770	
	(1.75)		(1.77)	
Observations	3770	680	3770†	680

Note: Absolute values of *t* statistics are in parentheses.
*Modeled with two-stage conditional logit model where the standard errors were corrected with bootstrapping.
†Significant at 5% (two-tailed test).
‡Significant at 1% (two-tailed test).

with using computers for health information. In the fixed-effects models, only Internet access was a significant instrument. Given that the HCP intervention provided more than just computerized health information, it may not be a valid instrument because it may be picking up other HCP effects. Yet if this instrument was excluded, the results were largely the same.

This study was limited to an analysis of self-reported doctor visits in the past year. Future research should consider other types of care as the effect of using computer health information may vary depending on the outcome. Offsetting changes in appropriate and inappropriate care would be a worthwhile effect, but not necessarily detectable with these data. For example, if after using information, a person sought care through an urgent care clinic instead of an emergency room, then this benefit would not be evident with our data. Similarly, we might detect no mean differences because users of health information may use more preventive care and less curative care, resulting in no change in the overall number of visits.

Other limitations of this study should also be considered. First, these data are from a household survey. As with any survey, there are potential self-reporting biases. We assumed that these biases were consistent across time and people, and if so then the study design minimizes these biases. However, idiosyncratic biases may still exist.

Many of the limitations of this study can be addressed by future research. A city in South Carolina has decided to replicate the HCP, without the media advertising. Will this town see similar results? In part, this depends on the generalizability of our findings, and in part, it depends on cultural differences in the use of the health information. With regard to the generalizability of our findings, we compared our sample with the Metropolitan Statistical Area (MSA) and national census data [35]. These analyses indicated that our sample is disproportionately female, has a higher proportion with some college education, and has more poor and wealthy households (results not shown). While this may raise questions about the generalizability of these findings, as the HCP is replicated elsewhere, we will gain a much greater insight to the effects of providing consumers with self-care information.

In conclusion, providing people with computerized health information is intuitively appealing. Some health plans may find this sufficient to justify investments in the health information infrastructure. Results from this paper show that using computerized health information is generally not associated with decreases in self-reported utilization, and, if anything, the use of computerized information is associated with slight increases in the probability of using health care. More research is needed to determine whether health information affects the health production function in other important ways, such as the location of care or the intensity of treatment.

We thank Merwyn Greenlick and Judith Hibbard for their comments. All errors are our own.

T.H.W. was supported in part by a dissertation grant from the Agency for Healthcare Research and Quality (R03 HS09997). H.B.J. was supported by the Robert Wood Johnson Foundation.

References

1 Modlin M. Vice President Gore launches free MEDLINE. NLM Newsline March–August 1997; 52(2–4).

2 DeNelsky SJ. The history of e-health. Paper presented at E-Health Connections: Health in the Digital Age; 2000 May 7–9; Stanford University.

3 Hibbard JH, Slovic P, Jewett JJ. Informing consumer decisions in health care: implications from decision making research. Milbank Q 1997;75: 395–414.

4 Fast J, Vosburgh RE, Frisbee WR. The effects of consumer education on consumer search. J Consumer Affairs 1989;23:65–90.

5 Jadad AR, Gagliardi A. Rating health information on the Internet: navigating to knowledge or to Babel? JAMA 1998;279:611–4.

6 Biermann JS, Golladay GJ, Greenfield ML, Baker LH. Evaluation of cancer information on the Internet. Cancer 1999;86:381–90.

7 Osman LM, Abdalla MI, Beattie JAG, et al. Reducing hospital admission through computer supported education for asthma patients. Grampian Asthma Study of Integrated Care (GRASSIC). BMJ 1994;308:568–71.

8 Lorig KR, Sobel DS, Stewart AL, et al. Evidence suggesting that a chronic disease self-management program can improve health status while reducing hospitalization—a randomized trial. Med Care 1999;37:5–14.

9 Piette JD, Weinberger M, McPhee SJ, et al. Do automated calls with nurse follow-up improve self-care and glycemic control among vulnerable patients with diabetes? Am J Med 2000;108:20–7.

10 Moore SH, LoGerfo J, Inui TS. Effect of a self-care book on physician visits: a randomized trial. JAMA 1980;243:2317–20.

11 Kemper DW. Self-care education: impact on HMO costs. Med Care 1982;20:710–8.

12 Carney N, Greenlick MR, Austin DF, et al. Assessment of Self-Care Manuals. Portland (OR): Oregon Health Sciences University; 2000.

13 Kenkel DS. Health behavior, health knowledge, and schooling. J Polit Econ 1991;99:287–305.

14 Hsieh CR, Lin SJ. Health information and the demand for preventive care among the elderly in Taiwan. Hum Resour 1997;32:308–33.

15 Borgmann A. Holding Onto Reality: The Nature of Information at the Turn of the Millennium. Chicago: University of Chicago Press; 1999.

16 Wagner TH, Hibbard JH. Who uses self-care books, advice nurses, and computers for health information? Int J Technol Assess Health Care 2001;17:590–600.
17 Wagner TH, Hibbard JH, Greenlick MR, Kunkel L. Does providing consumer health information affect self-reported medical utilization? Evidence from the healthwise communities project. Med Care 2001;39:836–47.
18 Wagner TH, Greenlick MR. When parents are given greater access to health information, does it affect pediatric utilization? Med Care 2001;39: 848–55.
19 Wagner TH, Hu TW, Hibbard JH. The demand for consumer health information. J Health Econ 2001;20:1059–75.
20 Grossman M. On the concept of health capital and the demand for health. J Polit Econ 1972;80:223–55.
21 Kenkel DS. The demand for preventive medical care. Appl Econ 1994;26:313–25.
22 Perneger TV, Etter JF, Rougemont A. Randomized trial of use of a monetary incentive and a reminder card to increase the response rate to a mailed health survey. Am J Epidemiol 1993;138:714–22.
23 Jones R, Pitt N. Health surveys in the workplace. comparison of postal, email, and World Wide Web methods. Occup Med (London) 1999;49:556–8.
24 Hogue C, Hargraves M, Collins K, editors. Minority Health in America: Findings and Policy Implications from the Commonwealth Fund Minority Health Survey. Baltimore: Johns Hopkins University Press; 2000.
25 Mullahy J. Much ado about two: reconsidering retransformation and the two-part model in health econometrics. J Health Econ 1998;17:247–81.
26 Kenkel D. Consumer health information and the demand for medical care. Rev Econ Stat 1990;72: 587–95.
27 Schafer JL. Analysis of Incomplete Multivariate Data. London: Chapman & Hall; 1997.
28 Amelia N, Honaker J, Joseph A, et al. A Program for Missing Data. 1998–98 ed. Boston: Harvard University; 1999.
29 King G, Honaker J, Joseph A, Scheve K. Analyzing incomplete political science data: an alternative algorithm for multiple imputation. Am Polit Sci Rev 2001;95:49–69.
30 Hausman JA, Taylor WE. Panel data and unobserved individual effects. Econometrica 1981;49: 1377–98.
31 Bound J, Jaeger DA, Baker RM. Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. J Am Stat Assoc 1995;90:443–50.
32 Staiger D, Stock JH. Instrumental variables regression with weak instruments. Econometrica 1997; 65:557–86.
33 Deaton A. The Analysis of Household Survey Data: A Microeconometric Approach to Development Policy. Baltimore: Johns Hopkins University Press; 1997.
34 Carlton J. Cheaper PCs start to attract new customers. Wall Street Journal, Monday, January 26, 1998; p. B1(E), B1(W) (col. 6).
35 US Census Bureau. Public Use Files from the 1990 Census. Washington (DC): US Department of Commerce; 1998.